Asymptotic Analysis of Distributed Multi-cell Beamforming

Subhash Lakshminaryana*†, Jakob Hoydis*†, Mérouane Debbah*† and Mohamad Assaad*
*Dept. of Telecommunications
SUPELEC, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France
Email: firstname.lastname@supelec.fr
†Alcatel-Lucent Chair on Flexible Radio

***Abstract*—We consider the problem of multi-cell downlink beamforming with N cells and K terminals per cell. Cooperation among base stations (BSs) has been found to increase the system throughput in a multi-cell set up by mitigating inter-cell interference. Most of the previous works assume that the BSs can exchange the instantaneous channel state information (CSI) of all their user terminals (UTs) via high speed backhaul links. However, this approach quickly becomes impractical as N and K grow large. In this work, we formulate a distributed beamforming algorithm in a multi-cell scenario under the assumption that the system dimensions are large. The design objective is the minimize the total transmit power across all BSs subject to satisfying the user SINR constraints while implementing the beamformers in a distributed manner. In our algorithm, the BSs would only need to exchange the channel statistics rather than the instantaneous CSI. We make use of tools from random matrix theory to formulate the distributed algorithm. The simulation results illustrate that our algorithm closely satisfies the target SINR constraints when the number of UTs per cell grows large, while implementing the beamforming vectors in a distributed manner.**

I. Introduction

Inter-cell interference mitigation has been an identified as an important consideration in the design of modern day cellular systems. Conventionally, the beamformer design problem in cellular systems was performed on a single cell basis treating the inter-cell interference as background noise [1]. However, such an approach is known to be leading to a suboptimal solution in the multi-cell context [2],[3]. Significant performance improvements may be obtained if the base stations (BSs) coordinate in jointly optimizing all of their beamformers at the same time, especially for the user terminals (UTs) at the cell edges. The intuition behind this is the following. Consider two UTs in adjacent cells close to their respective cell edges. The per-cell optimization solution would lead to the signal being steered in the direction of the respective edge UT by both of the BSs. It is easy to see that this solution would lead to a high level of interference. However, joint beamforming across the cells would lead to a solution in which the optimal tradeoff between signal enhancement and interference avoidance is found.

In recent years, a BS coordination technique known as network MIMO [3],[4] has been widely investigated to handle the inter-cell interference. The concept of network MIMO enables the BSs to share their CSI and user data through high-capacity backhaul links. If the BSs are allowed to cooperate without any restrictions on backhaul link capacity and processing delay, the multi-cell interference channel would be transformed into a broadcast channel in the downlink scenario without inter-cell interference. Many variants of network MIMO have been proposed considering the effects of limited backhaul capacity and imperfect CSI. A survey of the related results can be found in [5], [6]. However, sharing of CSI and the user data between the BSs demands high backhaul link capacity and computational power, which scale rapidly with the number of cells and the number of UTs, making this approach difficult to implement in a practical system.

Recently, distributed beamforming strategies that exploit only the locally available CSI have been developed where each BS balances the ratio between the signal gain at the intended terminal and the interference caused at other terminals [7],[8]. In these works, the design objective is to maximize the total sum rate of all the users subject to the antenna power constraints. They make use of the concept of virtual signal to interference and noise ratio (VSINR) approach which is shown to attain optimal rate points. However, the optimality of such algorithms have only been proved for the two user case [9]. To generalize the concept to the multi-user case, one needs to resort to heuristic strategies [10]. In [11], an outage minimizing power allocation has been proposed which requires a combination of perfect local CSI at the BSs and statistical CSI at a central processor.

Reference [12] provides an optimal algorithm for the multi-cell beamforming problem using convex optimization tools. Here, the design objective is to minimize the total transmitted power satisfying some SINR constraints of the UTs. The fundamental difference between this work and the previous mentioned works ([7]-[11]) is that in [12], each BS serves the UTs present in only its cell. Hence, the BSs share only the CSI of their UTs. There is no requirement of sharing the user data between the BSs. But, the solution provided in [12] cannot be implemented without the BSs exchanging the instantaneous CSIs of their UTs between each other. Sharing the CSI between the BSs calls for tremendous amount of data exchange between the BSs, specially in a fast fading scenario. A more important bottleneck is to share the CSI between the BSs under given delay constraints.

In this work, we extend the algorithm of [12] to be implemented in a distributed manner using tools from Random Matrix Theory. We provide a distributed beamforming algorithm in a multi-cell scenario when the dimension of the system becomes large (number of transmit antennas at the BS and the number of UTs in each cell). However, our algorithm provides good approximations even in the finite dimensional case. Our algorithm enables the BSs to compute the downlink beamforming vectors based on only their local CSI and the average channel statistics of the channels of other BSs. The BSs would only need to exchange the average channel statistics between themselves. This reduces the feedback load significantly since in a fast fading environment, the instantaneous channel realizations vary rapidly where as the channel statistics do not.

Throughout this work, we use boldface lowercase and uppercase letters to designate column vectors and matrices, respectively. For a matrix $\mathbf{X}$, $\mathbf{X}_{i,j}$ denotes the (i,j) entry of $\mathbf{X}$, and $\mathrm{tr}(\mathbf{X})$ denotes the trace of the matrix. $\mathbf{X}^T$ and $\mathbf{X}^H$ denote the transpose and complex conjugate transpose of matrix $\mathbf{X}$. We denote an identity matrix of size M as $\mathbf{I}_M$ and $\mathrm{diag}(x_1,...,x_M)$ is a diagonal matrix of size M with the elements x_i on its main diagonal. We use $\mathbf{x} \sim \mathcal{CN}(\mathbf{m},\mathbf{R})$ to state that the vector $\mathbf{x}$ has a complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$. We will use the notation $\xrightarrow{\text{a.s.}}$ to denote almost sure convergence.

II. Relevant results from Random Matrix Theory

In this section we give a brief overview of relevant results from random matrix theory which will be used in this work. These results are asymptotically exact when the dimensions of the cellular system grows infinitely large (in terms of the number of transmit antennas on each BS and the number of UTs) but provide already very good approximations for finite system dimensions.

We make extensive use of the following lemmas.

Lemma 1. *(Equation 2.2, [13]) Let $\mathbf{A}$ be a Hermitian invertible matrix of size $N \times N$, then for any vector $\mathbf{x} \in \mathbb{C}^N$ and scalar $\tau \in \mathbb{C}$ for which $\mathbf{A} + \tau\mathbf{x}\mathbf{x}^H$ is invertible,*

$$\mathbf{x}^H(\mathbf{A} + \tau\mathbf{x}\mathbf{x}^H)^{-1} = \frac{\mathbf{x}^H\mathbf{A}^{-1}}{1 + \tau\mathbf{x}^H\mathbf{A}^{-1}\mathbf{x}}$$

Lemma 2. *[14] Let $\mathbf{x} \in \mathbb{C}^N$, i.i.d. with zero mean, variance $1/N$, $\mathbf{A} \in \mathbb{C}^{N\times N}$ Hermitian with bounded spectral norm whose elements are independent of $\mathbf{x}$, then*

$$\mathbf{x}^H\mathbf{A}\mathbf{x} - \frac{1}{N}tr(\mathbf{A}) \xrightarrow{a.s.} 0$$

Lemma 3. *(Lemma 2.6, [13]) Let $z \in \mathbb{C}^+$ with $v = \mathrm{Im}(z)$ and $\mathbf{A}$ and $\mathbf{B}$ are $N \times N$ matrices with $\mathbf{B}$ being hermitian, $\tau \in \mathbb{R}$, and $\mathbf{q} \in \mathbb{C}^N$, then*

$$|tr\big((\mathbf{B} - z\mathbf{I})^{-1} - (\mathbf{B} + \tau\mathbf{q}\mathbf{q}^H - z\mathbf{I})^{-1})\mathbf{A}\big)| \leq \frac{||\mathbf{A}||}{v}$$

where $||\mathbf{A}||$ is the spectral norm of the matrix $\mathbf{A}$.

Next, we characterize the eigenvalue distribution of random matrices. For a hermitian $N \times N$ matrix $\mathbf{X}_N$, we denote the empirical spectral distribution (e.s.d.) by the notation $\mathbf{F}^{\mathbf{X}_N}$, which is defined for $x \in \mathbb{R}$ as

$$\mathbf{F}^{\mathbf{X}_N}(x) = \frac{1}{N}\sum_{j=1}^{N}\mathbf{1}_{\lambda_j \leq x}(x)$$

where $\{\lambda_1, \ldots, \lambda_N\}$ are the eigenvalues of $\mathbf{X}_N$ and $\mathbf{1}_{\lambda_j \leq x}$ is then indicator function whose value is equal to 1 if the eigenvalue λ_j is less than x, 0 otherwise. We denote the Stieltjes transform of the e.s.d of $\mathbf{X}_N$ by $m_{\mathbf{X}_N}(z)$ which is defined for z outside the support of $\mathbf{F}^{\mathbf{X}_N}$ as

$$m_{\mathbf{X}_N}(z) = \int_{-\infty}^{\infty}\frac{1}{\lambda - z}d\mathbf{F}^{\mathbf{X}_N}(\lambda) = \frac{1}{N}\mathrm{tr}(\mathbf{X}_N - z\mathbf{I}_N)^{-1}$$

The following theorem provides a deterministic equivalent of the Stieltjes transform of a Gram matrix $\mathbf{Y}_n\mathbf{Y}_n^H$ when $\mathbf{Y}_n$ has a variance profile.

Theorem 1. *(Theorem 2.4, [15]). Let $\mathbf{Y}_n$ be an $N \times n$ random matrix with independent elements having zero mean and variance profile $\mathbf{V}$ such that $\mathbb{E}[|(\mathbf{Y}_n)_{i,j}|^2] = \sigma^2_{i,j}$. Under some mild assumptions on the higher moments of the entries of $\mathbf{Y}_n$ and $\sigma^2_{i,j}, \forall i,j$ uniformly bounded from above, there exists a deterministic $N \times N$ matrix-valued function $\mathbf{\Psi}_n(z) = diag(\psi_1(z), \ldots, \psi_N(z))$, analytic in $\mathbb{C} - \mathbb{R}^+$ such that,*

$$\frac{1}{N}tr(\mathbf{Y}_n\mathbf{Y}_n^H - z\mathbf{I}_n)^{-1} - \frac{1}{N}tr(\mathbf{\Psi}_n(z)) \xrightarrow[a.s.]{n\to\infty,\frac{N}{n}\to c} 0, \quad \forall z \in \mathbb{C} - \mathbb{R}^+$$

whose elements are the unique solutions of the deterministic system of $N + n$ implicit equations

$$\psi_i(z) = \frac{-1}{z(1 + \frac{1}{n}\sum_{j=1}^{n}\sigma^2_{i,j}\tilde{\psi}_j(z))} \qquad \forall 1 \leq i \leq N, \quad (1)$$

$$\tilde{\psi}_j(z) = \frac{-1}{z(1 + \frac{1}{n}\sum_{i=1}^{N}\sigma^2_{i,j}\psi_i(z))} \qquad \forall 1 \leq j \leq n \quad (2)$$

such that $\frac{1}{N}tr(\mathbf{\Psi}_n(z))$ is the Stieltjes transform of a probability measure.

$\psi_i(z)$ and $\tilde{\psi}_j(z)$ can be obtained by intializing them to known values and iterating over Equations (1) and (2) until their values converge.

The differential of the Stieltjes transform of the matrix $\mathbf{Y}_n\mathbf{Y}_n^H$ can be calculated in the following way. Let us denote another deterministic matrix-valued function by the notation $\mathbf{\Psi}'_n(z) = \mathrm{diag}(\psi'_1(z), \ldots, \psi'_N(z))$. The elements of the matrix $\mathbf{\Psi}_n(z)$ can be evaluated using the deterministic system of $N + n$ equations,

$$\psi'_i(z) = \psi_i^2(z)\Big(\big(1 + \frac{1}{n}\sum_{j=1}^{n}\sigma^2_{i,j}\tilde{\psi}_j\big) + z\big(\frac{1}{n}\sum_{j=1}^{n}\sigma^2_{i,j}\tilde{\psi}'_j(z)\big)\Big) \quad (3)$$

$$\tilde{\psi}'_j(z) = \tilde{\psi}_j^2(z)\Big(\big(1 + \frac{1}{n}\sum_{i=1}^{N}\sigma^2_{i,j}\psi_i\big) + z\big(\frac{1}{n}\sum_{i=1}^{N}\sigma^2_{i,j}\psi'_i(z)\big)\Big) \quad (4)$$

where $\psi_i(z)$ and $\tilde{\psi}_j(z)$ are as defined in Theorem 1. Equations (3) and (4) can be obtained by differentiating (1) and (2) with respect to z, respectively. The derivative of the Stieltjes transform can then be calculated as $\frac{1}{N}\text{tr}\left(\mathbf{\Psi}_n^{'}(z)\right)$.

III. System Model and Algorithm Description

In this section, we present our system model and formulate the beamformer design as the solution to a convex optimization problem. The beamformer design problem consists of minimizing the total transmit power across all the BSs subject to SINR constraints of the UTs.

A. Sytem Model

We consider the problem of multi-cell beamforming across N cells and K UTs per cell where each BS is equipped with N_t antennas and and each UT has a single antenna. Each BS serves only the UTs in its cell. Let $\mathbf{h}_{i,j,k} \in \mathbb{C}^{N_t}$ denote the channel from the BS i to the k-th UT in cell j. We assume that the elements of the channel vector are Gaussian distributed, i.e., $\mathbf{h}_{i,j,k} \sim \mathcal{CN}(0, (\sigma_{i,j,k}^2/N_t)\mathbf{I}_{N_t})$, the variance of which depends upon the path loss model between BS i and UT (j,k). The channel variance has been scaled by the factor N_t to maintain the per antenna power constraint at each base station. Let $\mathbf{w}_{i,j} \in \mathbb{C}^{N_t}$ denote the transmit downlink beamforming vector for the j-th UT in cell i. Likewise, let $\Gamma_{i,j}$ denote the received SINR for the jth UT in cell i and $\gamma_{i,j}$ the corresponding target SINR. The received signal $y_{i,j} \in \mathbb{C}$ for the jth UT in cell i, is given by

$$y_{i,j} = \sum_{l=1}^{K} \mathbf{h}_{i,i,j}^H \mathbf{w}_{i,l} x_{i,l} + \sum_{m=1, m\neq i}^{N} \sum_{n=1}^{K} \mathbf{h}_{m,i,j}^H \mathbf{w}_{m,n} x_{m,n} + z_{i,j}$$

where $x_{i,j} \in \mathbb{C}$ represents the information signal for the j-th user in cell i and $z_{ij} \sim \mathcal{CN}(0,\sigma^2)$ is the corresponding additive white Gaussian complex noise. It has been shown in [12] that the sum power minimization problem can be cast into the following optimization problem given by

$$\begin{aligned} \min \quad & \sum_{i,j} \alpha_i \mathbf{w}_{i,j}^H \mathbf{w}_{i,j} \\ s.t. \quad & \Gamma_{i,j} \geq \gamma_{i,j}, \qquad i = 1 \dots N,\, j = 1 \dots K \end{aligned} \tag{5}$$

The power of the i^{th} BS is scaled by the factor α_i. The scaling factors α_i are assumed to be constant. Throughout this work, we assume that α_i is a precomputed constant. The received SINR in the downlink is given by the following expression

$$\Gamma_{i,j} = \frac{|\mathbf{w}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{l\neq j} |\mathbf{w}_{i,l}^H \mathbf{h}_{i,i,j}|^2 + \sum_{m\neq i,n} |\mathbf{w}_{m,n}^H \mathbf{h}_{m,i,j}|^2 + \sigma^2}$$

This problem has been solved in [12] using the uplink downlink duality approach. The authors formulate the corresponding dual uplink problem, given as

$$\begin{aligned} \max \quad & \sum_{i,j} \lambda_{i,j} \sigma^2 \\ s.t. \quad & \Lambda_{i,j} \geq \gamma_{i,j}, \qquad i = 1, \dots, N,\, j = 1, \dots, K \end{aligned} \tag{6}$$

where the uplink SINR, $\Lambda_{i,j}$ is given by

$$\Lambda_{i,j} = \frac{\lambda_{i,j} |\hat{\mathbf{w}}_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{(m,l)\neq(i,j)} \lambda_{m,l} |\hat{\mathbf{w}}_{i,j}^H \mathbf{h}_{i,m,l}|^2 + \alpha_i ||\hat{\mathbf{w}}_{i,j}||_2^2}$$

where $\hat{\mathbf{w}}$ denotes the corresponding uplink beamforming vector and $\lambda_{i,j}$ represents the dual variable associated with the optimization problem in (6). The $\lambda_{i,j}$ can be viewed as the dual uplink power.

B. Algorithm Design

We now provide a brief description of the beamforming algorithm presented in [12]. In what follows, without loss of generality, we set the noise variance σ^2 equal to 1.

Before introducing the algorithm, we define the following matrices. Let $\mathbf{H}_i = [\mathbf{h}_{i,1,1} \mathbf{h}_{i,1,2} \cdots \mathbf{h}_{i,m,n} \cdots \mathbf{h}_{i,N,K}]$ be the matrix whose columns are formed by the channel vectors from BS i to all the UTs across all the cells. Similarly, $\mathbf{\Lambda} = \text{diag}[\lambda_{1,1} \lambda_{1,2} \cdots \lambda_{m,n} \cdots \lambda_{N,K}]$ be a diagonal matrix with diagonal elements are the uplink power allocations. We also define the matrix $\mathbf{\Sigma}_i$ as

$$\mathbf{\Sigma}_i = \mathbf{H}_i \mathbf{\Lambda} \mathbf{H}_i^H \tag{7}$$

The details of the optimal uplink power allocation and the computation of uplink beamforming vectors provided in [12] are as follows.

- The optimal uplink power allocation $\lambda_{i,j}$ is evaluated using the iterative function

$$\lambda_{i,j} = \frac{1}{(1 + \frac{1}{\gamma_{i,j}}) \mathbf{h}_{i,i,j}^H (\mathbf{\Sigma}_i + \alpha_i \mathbf{I}_{N_t})_i^{-1} \mathbf{h}_{i,i,j}} \tag{8}$$

- The optimal receive uplink beamformers are given by

$$\hat{\mathbf{w}}_{i,j} = \Big(\sum_{m,l} \lambda_{m,l} \mathbf{h}_{i,m,l} \mathbf{h}_{i,m,l}^H + \alpha_i I \Big)^{-1} \mathbf{h}_{i,i,j} \tag{9}$$

- The optimal transmit downlink beamformers are given by

$$\mathbf{w}_{i,j} = \sqrt{\delta_{i,j}} \hat{\mathbf{w}}_{i,j} \tag{10}$$

The details of the calculation of the parameters $\delta_{i,j}$ are provided in [16].

However, as mentioned before, the solution provided in [12] cannot be implemented in a distributed manner. The computation of optimal uplink power allocations ($\lambda_{i,j}$) and the scaling factors ($\delta_{i,j}$) requires a central station which has the global system knowledge. We overcome this problem and formulate an algorithm which implements the beamforming problem in a distributed manner with the knowledge of only the channel statistics of the UTs under the assumption that the system dimensions grow large.

First note that the second term in the denominator of (8) can be approximated to the trace of the matrix $(\mathbf{\Sigma}_i + \alpha_i \mathbf{I}_{N_t})^{-1}$ using Lemma 2 when the matrix dimensions grow large. However, to apply the result of Lemma 2, the matrix $(\mathbf{\Sigma}_i + \alpha_i \mathbf{I}_{N_t})^{-1}$ should be made independent of the column

$\mathbf{h}_{i,i,j}$. Using Lemma 1 we can rewrite the second term in the denominator of (8) as

$$\mathbf{h}_{i,i,j}^H(\mathbf{\Sigma}_i + \alpha_i \mathbf{I}_{N_t})^{-1}\mathbf{h}_{i,i,j} = \Big(\frac{\mathbf{h}_{i,i,j}^H(\mathbf{\Sigma}_i^{'} + \alpha_i \mathbf{I}_{N_t})^{-1}\mathbf{h}_{i,i,j}}{1+\lambda_{i,j}\mathbf{h}_{i,i,j}^H(\mathbf{\Sigma}_i^{'} + \alpha_i \mathbf{I}_{N_t})^{-1}\mathbf{h}_{i,i,j}}\Big) \tag{11}$$

where $\mathbf{\Sigma}_i^{'}$ is the matrix $\mathbf{\Sigma}_i$ with the $\{i,i,j\}$th column removed. Note that the matrix $(\mathbf{\Sigma}_i^{'} + \alpha_i \mathbf{I}_{N_t})^{-1}$ is now independent of the entries of the column $\mathbf{h}_{i,i,j}$ and hence we can apply the the result of Lemma 2. Therefore, we have

$$\mathbf{h}_{i,i,j}^H(\mathbf{\Sigma}_i^{'} + \alpha_i \mathbf{I}_{N_t})^{-1}\mathbf{h}_{i,i,j} \xrightarrow[N_t\to\infty]{a.s.} \frac{\sigma_{i,i,j}^2}{N_t}\mathrm{tr}((\mathbf{\Sigma}_i^{'} + \alpha_i \mathbf{I}_{N_t})^{-1}) \tag{12}$$

when the dimensions become large. However, by application of the Rank-1 (Lemma 3) perturbation result, we have

$$\begin{aligned}\frac{1}{N_t}\mathrm{tr}((\mathbf{\Sigma}_i^{'} + \alpha_i \mathbf{I}_{N_t})^{-1}) &\approx \frac{1}{N_t}\mathrm{tr}((\mathbf{\Sigma}_i + \alpha_i \mathbf{I}_{N_t})^{-1}) \\ &= m_{\mathbf{\Sigma}_i}(-\alpha_i)\end{aligned} \tag{13}$$

Note that $m_{\mathbf{\Sigma}_i}(-\alpha_i)$ denotes the Stieltjes transform of the matrix $\mathbf{\Sigma}_i$ evaluated at the point $-\alpha_i$ which can be computed using Theorem 1. Hence,

$$\mathbf{h}_{i,i,j}^H(\mathbf{\Sigma}_i + \alpha_i \mathbf{I}_{N_t})^{-1}\mathbf{h}_{i,i,j} \xrightarrow[N_t\to\infty]{a.s.} \Big(\frac{\sigma_{i,i,j}^2 m_{\mathbf{\Sigma}_i}(-\alpha_i)}{1+\sigma_{i,i,j}^2\lambda_{i,j}m_{\mathbf{\Sigma}_i}(-\alpha_i)}\Big) \tag{14}$$

when the dimensions become large. We are now ready to present our distributed algorithm.

- The iterative function for evaluating the optimal uplink power allocation $\lambda_{i,j}$ is given by

$$\lambda_{i,j} = \Big((1+\frac{1}{\gamma_{i,j}})\big(\frac{\sigma_{i,i,j}^2 m_{\mathbf{\Sigma}_i}(-\alpha_i)}{1+\sigma_{i,i,j}^2\lambda_{i,j}m_{\mathbf{\Sigma}_i}(-\alpha_i)}\big)\Big)^{-1}$$

- The optimal receive uplink beamformers can be evaluated as in (9).
- The optimal transmit downlink beamformers are computed using

$$\mathbf{w}_{i,j} = \sqrt{\delta_{i,j}}\hat{\mathbf{w}}_{i,j}$$

Note that in our algorithm the iterative updates of the uplink parameter λ depend only on the channel statistics and not on the instantaneous CSIs. This results in a tremendous reduction of the amount of information to be exchanged between the BSs. In a typical fast fading channel, the channel statistics do not vary rapidly where as the channel realizations do. Also, for computing the uplink beamforming vectors, the BS i only needs to know the local channels between itself and different UTs.

We end the description of our algorithm by providing a distributed solution for the computation of δ based on only the channel statistics $\sigma_{i,j,k}$.

C. Computation of δ based upon the channel statistics

We use tools from random matrix theory once again to obtain an asymptotic approximation of δ, the scaling factor between the uplink and downlink beamforming vectors. In order to provide a clear explanation and to keep the notations simple, we explain our algorithm for the case of a single cell. The concept can be easily extended to the case of multiple cells.

Let us now consider a single cell with a BS having N_t antennas and K UTs in the cell. The channel from the BS to the UT i denoted by $\mathbf{h}_i \in \mathbb{C}^{N_t}$ is Gaussian distributed, $\mathbf{h}_i \sim \mathcal{CN}(0, (\sigma_i^2/N_t)\mathbf{I}_{N_t})$. Let $\mathbf{w}_i$ and $\hat{\mathbf{w}}_i$ be the transmit downlink and the receive uplink beamformers for the UT i, respectively. The uplink and the downlink beamformers are now related by the equation $\mathbf{w}_i = \sqrt{\delta_i}\hat{\mathbf{w}}_i$. The expression for the downlink SINR is then given by

$$\frac{|\mathbf{h}_i^H\mathbf{w}_i|^2}{\sum_{j\neq i}|\mathbf{h}_i^H\mathbf{w}_j|^2+\sigma^2} \geq \gamma_i \qquad i=1,\ldots,K \tag{15}$$

Without loss of generality, we assume $\sigma^2 = 1$. At the optimal point, all the SINR constraints must be active. Hence,

$$\frac{1}{\gamma_i}|\mathbf{w}_i^H\mathbf{h}_i|^2 = \sum_{j\neq i}|\mathbf{w}_j^H\mathbf{h}_i|^2 + 1 \qquad i=1,\ldots,K \tag{16}$$

Rewriting equation (16) in terms of the uplink beamformer vectors $\hat{\mathbf{w}}$, we have

$$\frac{\delta_i}{\gamma_i}|\hat{\mathbf{w}}_i^H\mathbf{h}_i|^2 = \sum_{j\neq i}\delta_j|\hat{\mathbf{w}}_j^H\mathbf{h}_i|^2 + 1 \qquad i=1,\ldots,K \tag{17}$$

Equation (17) provides us a set of linear equations in $\boldsymbol{\delta} = [\delta_1 \ldots \delta_K]$ which can be solved to obtain the values of $\boldsymbol{\delta}$. Let $\mathbf{G} \in \mathbb{R}^{K\times K}$ denote the matrix of coefficients of $\boldsymbol{\delta}$ with elements

$$\mathbf{G}_{i,j} \triangleq \begin{cases}\frac{1}{\gamma_i}|\hat{\mathbf{w}}_i^H\mathbf{h}_i|^2, & i=j\\ -|\hat{\mathbf{w}}_j^H\mathbf{h}_i|^2, & i\neq j\end{cases} \tag{18}$$

we can obtain $\boldsymbol{\delta}$ by solving

$$\boldsymbol{\delta} = \mathbf{G}^{-1}\mathbf{1}_K$$

where $\mathbf{1}_K$ denotes a $(K\times 1)$ vector with all elements equal to 1. However, it is easy to see that formulating the set of linear equations requires the global knowledge of all the channels. Henceforth, we seek for a deterministic approximation for the terms on the left and the right hand side of (16) based on only the channel statistics when the dimensions of the system grows large.

Let us first analyze the term on the LHS of equation (17). Note that we can write

$$|\hat{\mathbf{w}}_i^H\mathbf{h}_i|^2 = \mathbf{h}_i^H\hat{\mathbf{w}}_i\hat{\mathbf{w}}_i^H\mathbf{h}_i \tag{19}$$

Recall that the uplink beamforming vectors $\hat{\mathbf{w}}$ are constructed in a way similar to (9). Rewriting (9) in the single cell scenario,

$$\hat{\mathbf{w}}_i = (\mathbf{Q}+\alpha_i\mathbf{I})^{-1}\mathbf{h}_i$$

where the matrix $\mathbf{Q} = \sum_i \lambda_i \mathbf{h}_i \mathbf{h}_i^H$. Substituting for $\hat{\mathbf{w}}_i$ in (19), we get

$$\mathbf{h}_i^H \hat{\mathbf{w}}_i \hat{\mathbf{w}}_i^H \mathbf{h}_i = |\mathbf{h}_i^H (\mathbf{Q} + \alpha_i \mathbf{I})^{-1} \mathbf{h}_i|^2 \quad (20)$$

Let us now analyze term on the RHS of (20). Using arguments similar to the one used for deriving (14), we have

$$\mathbf{h}_i^H (\mathbf{Q} + \alpha_i \mathbf{I})^{-1} \mathbf{h}_i \xrightarrow[N_t \to \infty]{a.s.} \frac{\sigma_i^2 m_Q(-\alpha_i)}{1 + \lambda_i \sigma_i^2 m_Q(-\alpha_i)}$$

where $m_{\mathbf{Q}}(-\alpha_i)$ stands for the Stieltjes transform of the matrix $\mathbf{Q}$ evaluated at the point $-\alpha_i$. For notational convenience, we will introduce the notation

$$\eta_i = 1 + \sigma_i^2 \lambda_i m_{\mathbf{Q}}(-\alpha_i)$$

From (20) we have,

$$|\mathbf{w}_i^H \mathbf{h}_i|^2 \xrightarrow[N_t \to \infty]{a.s.} \left(\frac{\sigma_i^2 m_{\mathbf{Q}}(-\alpha_i)}{\eta_i} \right)^2$$

Now we analyze the terms on the RHS of equation (16). Rewriting the j^{th} term as

$$|\mathbf{w}_j^H \mathbf{h}_i|^2 = \mathbf{h}_i^H (\mathbf{Q} + \alpha_i \mathbf{I})^{-1} \mathbf{h}_j \mathbf{h}_j^H (\mathbf{Q} + \alpha_i \mathbf{I})^{-1} \mathbf{h}_i$$

we can follow a similar approach as before to find a deterministic approximation in the large system limit. Applying the Lemma 1 twice to remove the i^{th} and the j^{th} column from the matrix $(\mathbf{Q} + \alpha_i \mathbf{I})^{-1}$, and using the Rank-1 (Lemma 3) perturbation result, we can show that

$$|\mathbf{w}_j^H \mathbf{h}_i|^2 \xrightarrow[N_t \to \infty]{a.s.} \left(\frac{1}{N_t} \right) \frac{\sigma_i^2 \sigma_j^2 m'_{\mathbf{Q}}(-\alpha_i)}{\eta_i^2 \eta_j^2}$$

where $m'_{\mathbf{Q}}(-\alpha_i)$ is the differential of $m_{\mathbf{Q}}(z)$ with respect to z evaluated at the point $-\alpha_i$ which can be computed using the equations (3) and (4).

Hence we formulated the deterministic approximation for the terms of equation (16). We now have deterministic approximations for all the elements of the matrix $\mathbf{G}$.

$$\mathbf{G}_{i,j} \xrightarrow[N_t \to \infty]{a.s.} \begin{cases} \left(\frac{\sigma_i^2 m_{\mathbf{Q}}(-\alpha_i)}{\eta_i} \right)^2, & i = j \\ -\frac{1}{N_t} \frac{\sigma_i^2 \sigma_j^2 m'_{\mathbf{Q}}(-\alpha_i)}{\eta_i^2 \eta_j^2}, & i \neq j \end{cases}$$

The algorithm for the computation of $\boldsymbol{\delta}$ can be easily extended to a multi-cell scenario by taking care of the terms representing the out of cell interference.

IV. Simulation Results

In this section, we verify the accuracy of our distributed beamforming algorithm in satisfying the UT SINR constraints. We consider a hexagonal cellular system with a cluster of 3 cells as shown in Figure 1. Each cell has a BS which serves only the UTs in its cell. The number of transmit antennas on each BS scales with the number of UTs in the cell such that their ratio is a finite constant. In particular, we assume $N_t = K$.

We consider a distance dependent path loss model in which the UTs are assumed to be arbitrarily scattered inside each cell. In this case, the path loss factor from UT k in cell i to BS j is given as

$$\sigma_{i,j,k}^2 = \left(\frac{1}{d_{i,j,k}} \right)^\beta$$

where $d_{i,j,k}$ is the distance between UT k in cell i to BS j, normalized to the maximum distance within a cell, and β is the path loss exponent which lies usually in the range from 2 to 5 dependent on the radio environment. We normalize the variance of the total received noise to $\sigma^2 = 1$. We also assume that no user terminal is within a normalized distance of 0.1 from the closest BS.

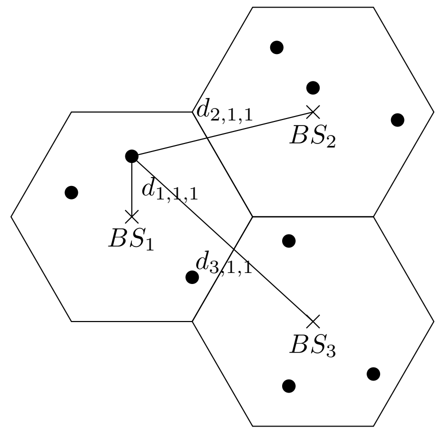

Fig. 1. Example of a Network with 3 Cells. The crosses represent the location of the BSs and the dots represent the location of the UTs randomly scattered inside the cells. The distances of a UT from the three BSs are also provided.

We run our simulations for the asymptotic beamforming algorithm for T different channel realizations. The outage metric we use in our simulations is the normalized mean squared error (NMSE) of the received SINR, averaged over the total number of channel realizations and UTs in all cells. More specifically, let us denote the received SINR for user j in cell i at the t^{th} channel realization by $\Gamma_{i,j}^t$. The outage metric denoted by ϵ is calculated as follows

$$\epsilon = \frac{1}{TNK} \sum_{t,i,j} \frac{(\Gamma_{i,j}^t - \gamma_{i,j})^2}{\gamma_{i,j}^2}$$

Figure 2 shows the plot of our outage metric against the number of UTs for target SINRs of 0dB and 5dB for all the users. It can be seen that the outage metric ϵ asymptotically approaches zero as the number of UTs grow large. The result can be interpreted as follows. As the system dimensions become large, the SINR constraints are satisfied for the UTs for every given channel realization making our algorithm asymptotically optimal.

V. Conclusion

In this work, we have derived a distributed beamforming algorithm using only locally available CSI at the BS and some statistical side information of the channel gains to other

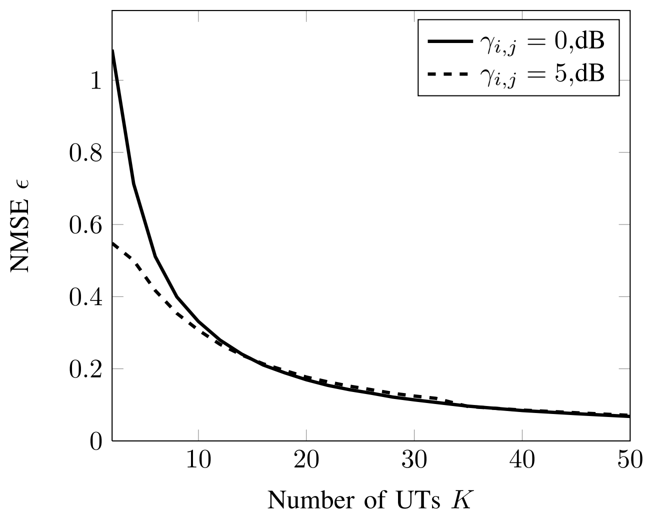

Fig. 2. NMSE ϵ of the SINR averaged over 1000 different channels realizations as a function of the number of UTs for the path loss exponent $\beta = 3.6$ and given target SINR.

UTs, when the system dimensions become large. Compared to the existing works, our algorithm incurs the least burden in terms of the information exchange between BSs and is asymptotically optimal. The future direction of this work would be to formulate an optimum algorithm for a smaller number of UTs in each cell taking the statistical distribution of the SINR at the UTs into account.

REFERENCES

[1] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO Paradigm," *Signal Processing Magazine, IEEE*, vol. 24, no. 5, pp. 36 –46, sept. 2007.

[2] S. Shamai and B. Zaidel, "Enhancing the Cellular Downlink Capacity via Co-Processing at the Transmitting End," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 3, 2001, pp. 1745 –1749 vol.3.

[3] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *Wireless Communications, IEEE*, vol. 13, no. 4, pp. 56 –61, aug. 2006.

[4] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming Intercell Interference in Indoor Wireless Systems," in *Proc. Asilomar Conference on Signals, Systems and Computers (ACSSC'07)*, 2007, pp. 83–87.

[5] S. Shamai, O. Somekh, O. Simeone, A. Sanderovich, B. M. Zaidel, and H. V. Poor, "Cooperative Multi-Cell Networks: Impact of Limited-Capacity Backhaul and Inter-Users Links," in *Proc. Joint Workshop on Coding and Communications (JWCC'07)*, Durnstein, Austria, Oct. 2007, pp. 14–16.

[6] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multi-cell MIMO Cooperative Networks: A New Look at Interference," *IEEE J. Sel. Areas Commun.*, 2010, to appear.

[7] R. Zakhour and D. Gesbert, "Coordination on the MISO Interference Channel Using the Virtual SINR Framework," in *Proc. International ITG Workshop on Smart Antennas*, 2009.

[8] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Distributed Multicell and Multiantenna Precoding: Characterization and Performance Evaluation," in *Proc. Global Communications Conference (GLOBECOM'09)*, Dec. 2009.

[9] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete Characterization of the Pareto Boundary for the MISO Interference Channel," *IEEE Trans. Signal Process.*, vol. 56, no. 10-2, pp. 5292–5296, 2008.

[10] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative Multicell Precoding: Rate Region Characterization and Distributed Strategies with Instantaneous and Statistical CSI," *IEEE Trans. Signal Process.*, Dec. 2009, submitted. [Online]. Available: http://www.eurecom.fr/~gesbert/papers/TSP09_emil.pdf

[11] M. Kobayashi, M. Debbah, and J.-C. Belfiore, "Outage Efficient Strategies for Network MIMO with Partial CSIT," in *Proc. International Symposium on Information Theory (ISIT'09)*, Jul. 2009.

[12] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multi-antenna wireless system," in *42nd Annual Conference on Information Sciences and Systems, CISS 2008, Princeton, NJ, USA, 19-21*.

[13] J. W. Silverstein and Z. Bai, "On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices," *Journal of Multivariate Analysis*, 1995.

[14] Z. Bai and J. W. Silverstein, "No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices," *Annals of Probability*, vol. 26, pp. 316–345, 1998.

[15] W. Hachem, P. Loubaton, and J. Najim, "Deterministic Equivalents for Certain Functionals of Large Random Matrices," *Annals of Applied Probability*, vol. 17, p. 875, 2007.

[16] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6-1, pp. 2646–2660, 2007.